

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-238142

(43)Date of publication of application : 09.09.1997

(51)Int.Cl. H04L 12/28
H04L 12/46
H04L 12/56
H04Q 3/00

(21)Application number : 08-043542

(71)Applicant : TOSHIBA CORP

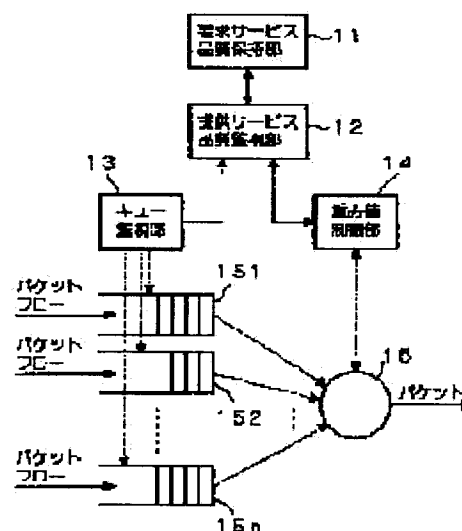
(22)Date of filing : 29.02.1996

(72)Inventor : YAMATO KATSUMI

(54) METHOD AND DEVICE FOR ASSIGNING NETWORK RESOURCE**(57)Abstract:**

PROBLEM TO BE SOLVED: To surely provide service quality on requirements by monitoring a difference between the service quality and the quality of requirement received from an assigned resource and revising the assignment of resources so as to decrease the difference.

SOLUTION: A packet transfer device holds packets for each logically or physically different flow into buffer sections 151, 152,... as queues tentatively. A transfer processing section 16 selects a queue based on a weight assigned to each flow and sends a packet. A requirement service quality holding section 11 stores quality items such as a transfer speed, a transfer delay and its request level corresponding to an identifier of a packet flow. A provided service quality monitor section 12 calculates the quality of service served at present based on the length of the queue from the queue monitor section 13 and weight data from a weight control section 14. The weight control section 14 revises the weight assigned to each flow, based on the difference between the calculated present served quality level and the requirement level.

**LEGAL STATUS**

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision]

of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

* NOTICES *

JPO and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] It is the network resource quota approach which assigns a network resource with the packet transfer equipment which transmits the packet which is one bundle of unit by which communication link information was subdivided. The network resource which can be offered for a packet transfer for every group of the flow classified based on the packet identifier currently described in each packet, or a flow The specified quantity [every] logic target which uses it in order to transmit the packet belonging to the group of this flow or a flow is assigned. The quality of service which the group of this flow or a flow wears with the amount of network resources assigned now about the group of each flow or a flow, So that the group of this flow or a flow may ask for the difference between the qualities of service demanded beforehand and may decrease this difference in the group of each flow or a flow The network resource quota approach characterized by changing the assignment of the amount of network resources to the group of each flow or a flow.

[Claim 2] Modification of the amount of network resources to the group of each of said flow or a flow is the network resource quota approach according to claim 1 characterized by to be that to which the quality of service which the quality of service covered now increases the amount of network resources to the group of the flow with which are not satisfied of the quality of service demanded beforehand, or a flow, and has covered now decreases the amount of network resources to the group of the flow with which are satisfied of the quality of service demanded beforehand, or a flow.

[Claim 3] Augend to the group of the flow to which the amount of network resources is made to increase, or a flow is made into the homogeneity value a in all the groups of this flow or a flow. The decrement to the group of the flow which decreases the amount of network resources, or a flow is made into the homogeneity value b in all the groups of this flow or a flow. When the number of the groups of the flow to which a network resource is made to increase, or a flow is set to m1 and the number of the groups of the flow which decreases a network resource, or a flow is set to m2 The network resource quota approach according to claim 2 characterized by choosing said homogeneity value a and said homogeneity value b so that $axm1 = bxm2$ may be materialized.

[Claim 4] The quality of service which the group x of said flow or a flow wears with the amount of network resources assigned now is set to Q_x . The group x of said flow or a flow sets to R_x the quality of service demanded beforehand. Augend to the group j of the flow to which the amount of network resources is made to increase, or a flow is made into $c_x (R_j - Q_j)$ (c is a homogeneity value). The decrement to the group k of the flow which decreases the amount of network resources, or a flow is made into $d_x (Q_k - R_k)$ (d is a homogeneity value). It is \sum_j about the total about the group of the flow to which said amount of network resources is made to increase, or a flow. It carries out. It is \sum_k about the total about the group of the flow which decreases said amount of network resources, or a flow. When it carries out, it is $c_x \sum_j (R_j - Q_j) - d_x \sum_k (Q_k - R_k)$ about said homogeneity value c and said homogeneity value d. The network resource quota approach according to claim 2 characterized by choosing so that it may *****.

[Claim 5] It is the network resource quota approach which assigns a network resource with the packet transfer equipment which transmits the packet which is one bundle of unit by which communication link information was subdivided. The network resource which can be offered for a packet transfer for every group of the flow classified based on the packet identifier currently described in each packet, or a flow The specified quantity [every] logic target which uses it in order to transmit the packet belonging to the group of this flow or a flow is assigned. The relation of the quality of service worn with the amount of network resources assigned now among the groups of each flow which is demanding the quality of service for the information which shows relative superiority or inferiority, or a flow of dominance, When what the difference has produced exists between the relation of the quality of service demanded beforehand of dominance The network resource quota approach characterized by changing the assignment of the amount of network resources to the group of the corresponding flow or a flow so that the relation of the quality of service demanded beforehand of dominance may be realized.

[Claim 6] It is that for which the value which shows logically the level of the service which the group of said flow which requires this quality of service, or a flow should receive as said quality of service is used. Modification of the amount of network resources to the group of said each flow or a flow In the group of each flow belonging to level which asks for the quality of service worn with the amount of network resources assigned now, and is mutually different about the group of each flow belonging to each level, or a flow, or a flow Among the qualities of service which the group of the flow belonging to the first level which requires the more excellent quality of service, or a flow wears, at least one quality When inferior compared with the quality of service which either of the groups of the flow belonging to the second level which requires the quality of service inferior to this first level, or a flow wears The amount of network resources currently assigned to the group of the flow belonging to this first level or a flow is increased. The network resource quota approach according to claim 5 characterized by being what decreases the amount of network resources currently assigned to the group of the flow belonging to this second level, or a flow.

[Claim 7] Augend to the group of the flow belonging to said first level to which the amount of network resources is made to increase, or a flow is made into the homogeneity value alpha in all the groups of this flow or a flow. The decrement to the group of the flow belonging to said second level which decreases the amount of network resources, or a flow is made into the homogeneity value beta in all the groups of this flow or a flow. When the number of the groups of the flow to which a network resource is made to increase, or a flow is set to n_1 and the number of the groups of the flow which decreases a network resource, or a flow is set to n_2 The network resource quota approach according to claim 6 characterized by choosing said homogeneity value alpha and said homogeneity value beta so that $\alpha n_1 = \beta n_2$ may be materialized.

[Claim 8] The quality of service which the group x of said flow or a flow wears with the amount of network resources assigned now is set to Q_x . The worst quality of service is set to Q_{1min} among the qualities of service worn in the group of the flow belonging to said first level, or a flow. The best quality of service is set to Q_{2max} among the qualities of service worn in the group of the flow belonging to said second level, or a flow. Augend to the group s of the flow belonging to said first level to which the amount of network resources is made to increase, or a flow is made into $\gamma_{max} (Q_{2max} - Q_s)$ (γ is a homogeneity value). The decrement to the group t of the flow belonging to said second level which decreases the amount of network resources, or a flow is made into $\delta_{max} (Q_t - Q_{1min})$ (δ is a homogeneity value). It is σ_{max} about the total about the group of the flow to which the amount of network resources is made to increase, or a flow. It carries out. It is σ_{min} about the total about the group of the flow which decreases the amount of network resources, or a flow. When it carries out, it is $\gamma_{max} \sigma_{max} s = (Q_{2max} - Q_s) \delta_{max} \sigma_{min} t (Q_t - Q_{1min})$ about said homogeneity value γ and said homogeneity value δ .

The network resource quota approach according to claim 6 characterized by choosing so that it may *****.

[Claim 9] The network resource which can be offered for a transfer of the packet which is one

bundle of unit by which communication link information was subdivided The means assigned to the specified quantity [every] logic target which uses it in order to transmit the packet belonging to the group of this flow or a flow for every group of the flow classified based on the packet identifier currently described in each packet, or a flow, A means by which the group of each flow or a flow holds the quality of service demanded beforehand or its relative relation of dominance, It has a means to supervise whether the quality of service which the group of this flow or a flow wears with the amount of network resources assigned now about the group of this flow or a flow, or its relative relation of dominance is what satisfies the contents of maintenance of said means to hold. Said means to assign is network resource quota equipment characterized by being what changes assignment of the amount of network resources based on the monitor result of said means to supervise.

[Translation done.]

* NOTICES *

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.*** shows the word which can not be translated.

3.In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates to the network resource quota approach and equipment which assign logically the network resource secured with packet transfer equipment per group of a flow unit or a flow.

[0002]

[Description of the Prior Art] ATM (Asynchronous Transfer Mode; Asynchronous Transfer Mode) When the demand of the quality of service to a packet transfer is made from the user who transmits and receives this packet on the occasion of a transfer of the packet (variable length) defined in IP (Internet Protocol) level used in case the communication link using the fixed-length packet called the cel in a communication network or the Internet is performed, a communication network is wanted to offer a transfer of a packet which satisfies such demand quality. Here, the relative quality at the time of comparing with the absolute quality of the transfer rate, and the max / average transfer delay which the packet of arbitration wears, or the quality which other packets called a service level wear is expressed as a quality of service.

[0003] In order to offer these demand qualities of service, some scheduling algorithms of the packet transfer which becomes indispensable in a communication network are examined. For example, it sets in the packet transfer section holding the packet buffer according to the queuing model based on the policy called FIFO (First In First Out) transmitted from the old packet of the time of arrival. Priority is beforehand assigned to every [to which the packet to which this packet transfer section gives its service belongs] flow (depending on the case, it is the group of a flow). There is the packet transmittal mode of transmitting from the packet which belongs to the high flow of priority regardless of the time of arrival (for example, L.Kleinrock:"Queueing Systems, Volume II", and Wiley-Interscience). Moreover, each packet which performed weighting to each flow and reached the packet transfer section The WFQ (Weighted Fair Queueing) method that the call forwarding service frequency worn in proportion to the weight assigned to the flow to which this packet belongs is determined is also considered. It Clark(s). for example, D.D. — Others: "SupportingReal-Time Applications in an Integrated Services Packet Network : Architecture and Mechanism", ACM SIGCOMM'92. Since the band equivalent to the above-mentioned weight is assigned to each flow fixed by the WFQ method, it is possible to offer service which can be offered in this band to the packet belonging to each flow.

[0004]

[Problem(s) to be Solved by the Invention] With the packet transmittal mode based on the priority mentioned above, it is thought that the opportunity of a packet transfer needs to add the control for not being given rarely, consequently performing the QA to the low flow of priority to the above-mentioned priority control, and needs to offer it to the low flow (or group of a flow) of priority. Moreover, with the packet transmittal mode based on above WFQ, it is thought by differing from the quality of service which the quality of service worn according to the weight assigned to each flow expected at the beginning, consequently coping with such an event flexibly that dynamic control which always fills a demand quality of service is required.

[0005] This invention is made in consideration of the above-mentioned situation, and aims at

offering the network resource quota approach and equipment for offering certainly the quality of service which the group of each flow or a flow requires to the group of all flows or flows.

[0006]

[Means for Solving the Problem] The packet whose this invention (claim 1) is one bundle of unit by which communication link information was subdivided With the packet transfer equipment which transmits for example, (an ATM cel and an IP packet), a network resource It is the network resource quota approach which assigns (for example, a packet transfer rate). The network resource which can be offered for a packet transfer for every group of the flow classified based on the packet identifier currently described in each packet, or a flow The specified quantity [every] logic target which uses it in order to transmit the packet belonging to the group of this flow or a flow is assigned. The quality of service which the group (x) of this flow or a flow wears with the amount of network resources assigned now about the group of each flow or a flow (Qx), So that the group (x) of this flow or a flow may ask for the difference between the qualities of service (Rx) demanded beforehand and may decrease this difference ($|R_x - Q_x|$) in the group of each flow or a flow It is characterized by changing the assignment of the amount of network resources to the group of each flow or a flow.

[0007] It is desirable to make dynamically a change of the assignment of the amount of network resources to the group of each above-mentioned flow or a flow repeatedly to predetermined timing. The group of VPI/VCI (a virtual path / channel identifier) for example, in an ATM communication mode or the flow label in IP protocol (version IPv6), and a source address etc. is used for a packet identifier.

[0008] the group of a flow — for example, two or more flows which require the same quality of service are adjusted as one group, and a network resource is assigned to the group of a flow in this case.

[0009] According to this invention, it becomes possible to offer a packet transfer to which the quality of service which each requires is satisfied certainly to the group of each flow which uses absolute quality, such as a transfer rate, or max / average transfer delay, as a quality of service, or a flow.

[0010] It is characterized by for this invention (claim 2) to be what decreases the amount of network resources to the group of the flow with which the quality of service which modification of the amount [on claim 1 and as opposed to the group of each of said flow or a flow] of network resources increases the amount of network resources to the group of the flow with which the quality of service covered now is not satisfied of the quality of service demanded beforehand, or a flow, and has been covered now is satisfied of the quality of service demanded beforehand, or a flow.

[0011] This invention (claim 3) makes augend to the group of the flow to which the amount of network resources is made to increase, or a flow the homogeneity value a in all the groups of this flow or a flow in claim 2. The decrement to the group of the flow which decreases the amount of network resources, or a flow is made into the homogeneity value b in all the groups of this flow or a flow. When the number of the groups of the flow to which a network resource is made to increase, or a flow is set to m1 and the number of the groups of the flow which decreases a network resource, or a flow is set to m2, it is characterized by choosing said homogeneity value a and said homogeneity value b so that $axm1 = bxm2$ may be materialized.

[0012] This invention (claim 4) sets to Qx the quality of service which the group x of said flow or a flow wears with the amount of network resources assigned now in claim 2. The group x of said flow or a flow sets to Rx the quality of service demanded beforehand. Augend to the group j of the flow to which the amount of network resources is made to increase, or a flow is made into cx ($R_j - Q_j$) (c is a homogeneity value). The decrement to the group k of the flow which decreases the amount of network resources, or a flow is made into dx ($Q_k - R_k$) (d is a homogeneity value). It is \sum_j about the total about the group of the flow to which said amount of network resources is made to increase, or a flow. It carries out. It is \sum_k about the total about the group of the flow which decreases said amount of network resources, or a flow. When it carries out, it is characterized by choosing said homogeneity value c and said homogeneity value d so that $c \sum_j (R_j - Q_j) = d \sum_k (Q_k - R_k)$ may be materialized.

[0013] This invention (claim 5) is the network resource quota approach which assigns a network resource with the packet transfer equipment which transmits the packet which is one bundle of unit by which communication link information was subdivided. The network resource which can be offered for a packet transfer for every group of the flow classified based on the packet identifier currently described in each packet, or a flow. The specified quantity [every] logic target which uses it in order to transmit the packet belonging to the group of this flow or a flow is assigned. The relation of the quality of service worn with the amount of network resources assigned now among the groups of each flow which is demanding the quality of service for the information which shows relative superiority or inferiority, or a flow of dominance, When what the difference has produced exists between the relation of the quality of service demanded beforehand of dominance. It is characterized by changing the assignment of the amount of network resources to the group of the corresponding flow or a flow so that the relation of the quality of service demanded beforehand of dominance may be realized.

[0014] It is desirable to make dynamically a change of the assignment of the amount of network resources to the group of each above-mentioned flow or a flow repeatedly to predetermined timing. According to this invention, it becomes possible to offer a packet transfer to which the quality of service which each requires is satisfied certainly to the group of each flow which requires a relative quality of service like a service level, or a flow.

[0015] This invention (claim 6) is set to claim 5. As said quality of service. It is that for which the value which shows logically the level of the service which the group of said flow which requires this quality of service, or a flow should receive is used. Modification of the amount of network resources to the group of said each flow or a flow. In the group of each flow belonging to level which asks for the quality of service worn with the amount of network resources assigned now, and is mutually different about the group of each flow belonging to each level, or a flow, or a flow. Among the qualities of service which the group of the flow belonging to the first level which requires the more excellent quality of service, or a flow wears, at least one quality. When inferior compared with the quality of service which either of the groups of the flow belonging to the second level which requires the quality of service inferior to this first level, or a flow wears. The amount of network resources currently assigned to the group of the flow belonging to this first level or a flow is increased, and it is characterized by being what decreases the amount of network resources currently assigned to the group of the flow belonging to this second level, or a flow.

[0016] This invention (claim 7) makes augend to the group of the flow which belongs to said first level to which the amount of network resources is made to increase in claim 6, or a flow. The homogeneity value α in all the groups of this flow or a flow. The decrement to the group of the flow belonging to said second level which decreases the amount of network resources, or a flow is made into the homogeneity value β in all the groups of this flow or a flow. When the number of the groups of the flow to which a network resource is made to increase, or a flow is set to n_1 and the number of the groups of the flow which decreases a network resource, or a flow is set to n_2 , it is characterized by choosing said homogeneity value α and said homogeneity value β so that $\alpha n_1 = \beta n_2$ may be materialized.

[0017] This invention (claim 8) sets to Q_x the quality of service which the group x of said flow or a flow wears with the amount of network resources assigned now in claim 6. The worst quality of service is set to Q_{1min} among the qualities of service worn in the group of the flow belonging to said first level, or a flow. The best quality of service is set to Q_{2max} among the qualities of service worn in the group of the flow belonging to said second level, or a flow. Augend to the group s of the flow belonging to said first level to which the amount of network resources is made to increase, or a flow is made into γ_{max} ($Q_{2max} - Q_s$) (γ is a homogeneity value). The decrement to the group t of the flow belonging to said second level which decreases the amount of network resources, or a flow is made into δ_{t} ($Q_t - Q_{1min}$) (δ is a homogeneity value). It is σ_{total} about the total about the group of the flow to which the amount of network resources is made to increase, or a flow. It carries out. It is σ_{total} about the total about the group of the flow which decreases the amount of network resources, or a flow. When it carries out, it is characterized by choosing said homogeneity value γ and said homogeneity value

delta so that $\gamma_{\max} \sigma_s = (Q_2 \max - Q_s) \Delta \sigma_t (Q_t - Q_1 \min)$ may be materialized.

[0018] This invention (claim 9) the network resource which can be offered for a transfer of the packet which is one bundle of unit by which communication link information was subdivided The means assigned to the specified quantity [every] logic target which uses it in order to transmit the packet belonging to the group of this flow or a flow for every group of the flow classified based on the packet identifier currently described in each packet, or a flow, A means by which the group of each flow or a flow holds the quality of service demanded beforehand or its relative relation of dominance, It has a means to supervise whether the quality of service which the group of this flow or a flow wears with the amount of network resources assigned now about the group of this flow or a flow, or its relative relation of dominance is what satisfies the contents of maintenance of said means to hold. Said means to assign is network resource quota equipment characterized by being what changes assignment of the amount of network resources based on the monitor result of said means to supervise.

[0019] It is desirable to make dynamically a change of the assignment of the amount of network resources to the group of each above-mentioned flow or a flow repeatedly to predetermined timing. According to this invention, it becomes possible to offer a packet transfer to which the quality of service which each requires is satisfied certainly to the group of each flow which requires a relative quality of service like the group of each flow which uses absolute quality, such as a transfer rate, or max / average transfer delay, as a quality of service, or a flow, and a service level, or a flow.

[0020]

[Embodiment of the Invention] Hereafter, the gestalt of implementation of invention is explained, referring to a drawing. The packet transfer section prepared in the packet transfer equipment concerning this operation gestalt has a configuration as shown in drawing 1 . In the packet transfer section which performs scheduling of two or more packet flow sending out, the packet which arrived is held temporarily within 151-15n (queue) of queues different logically or physically for every flow to which this packet belongs, and in the transfer processing section 16, the flow (queue) to which the packet which should be transmitted belongs is chosen based on the "weight" currently assigned to each flow, and it sends out to it behind.

[0021] Here, the variable-length packet as which the format etc. is beforehand determined in protocols, such as a fixed-length packet which it is one bundle of unit which subdivided this information in case a packet transmits communication link information, for example, is called the cel in an ATM communication mode, and TCP, IP, is shown. Moreover, the packet identifier currently described in said packet expresses the set of the same packet, or the set of the packet which has one of identifiers among two or more packet identifiers specified beforehand as a flow, and in an ATM communication mode, if a version is IPv6 in the case of VPI/VCI (a virtual path / channel identifier) and an IP packet, both the values of a flow label and a source address will prescribe a packet identifier to a meaning.

[0022] In the transfer processing section 16, scheduling about the packet transfer between two or more flows is performed so that the value to which the amount of network resources (packet transfer rate) given to each flow is proportional to the value of the weight beforehand assigned to the flow may be taken. In addition, this scheduling function is realizable by adopting for example, a WFQ method.

[0023] The method which adopts the amount of network resources assigned to a corresponding flow as a weight value used with this operation gestalt, or the method which adopts the rate that this flow can use an output circuit can be considered. In the case of the former, total of a weight value serves as the amount of network resources which can offer an output circuit, and, in the case of the latter, total of a weight value is set to 1.

[0024] The packet transfer section of this operation gestalt holds the weight value control section 14 which changes dynamically the weight value which controls the amount of network resources assigned to each flow. This supervises whether at the offer quality-of-service Monitoring Department 12, satisfying service is made to all flows based on the information in the queue Monitoring Department 13 which supervises the number of are recording packets for every flow, and the present weight value, and changes the weight value of the flow of arbitration

dynamically if needed. For example, a weight value is changed, when satisfaction is not provided with demand qualities of service which each flow required, such as packet transfer delay and a packet transfer rate, or when the group of the flow from which the relation of a service level of dominance and the relation of the quality of service currently offered of dominance differ exists. [0025] In addition, to mention later in detail, these can be treated as "a group of a flow" about some flows, for example, all the flows that require the same service level, a weight value can be assigned for every group of a flow, and a packet transfer can also be controlled.

[0026] Hereafter, the operation gestalt of this invention is further explained to a detail. In addition, below, it explains as what assigns a weight value for every flow. The packet transfer section prepared in the packet transfer equipment applied to this operation gestalt at drawing 1 is shown.

[0027] This packet transfer section is equipped with the demand quality-of-service attaching part 11, the offer quality-of-service Monitoring Department 12, the queue Monitoring Department 13, the weight value control section 14, the buffer sections 151-15n, and the transfer processing section 16.

[0028] The demand quality-of-service attaching part 11 holds the demand quality of service to each packet flow. The offer quality-of-service Monitoring Department 12 computes the quality of service currently offered to each packet flow in this packet transfer section, and compares this quality of service with the demand quality of service currently held within the demand quality-of-service attaching part 11.

[0029] The queue Monitoring Department 13 supervises the number of are recording packets for every packet flow. The weight value control section 14 changes the "weight value" which determines the transfer rate (band) assigned to each packet flow according to directions of the offer quality-of-service Monitoring Department 12.

[0030] The buffer sections 151-15n hold each packet flow. The transfer processing section 16 extracts and transmits the packet of the arbitration in 151-15n of buffer sections.

[0031] As the former is shown in drawing 7 here, in the demand quality-of-service attaching part 111 which received the demand of the quality of service from each flow, the weight value assigned to each flow is determined fixed, or the priority control based on a demand quality of service is performed. On the other hand, it is made to change dynamically based on the difference of the quality of service as which each flow required this weight value, and the actually covered quality of service, and enables it to offer certainly the quality of service which each flow requires from all flows with this operation gestalt.

[0032] Hereafter, it explains in more detail about the configuration and actuation of this packet transfer section. The demand quality-of-service attaching part 11 holds information as shown in drawing 2. As maintenance information in the demand quality-of-service attaching part 11, the item and demand quality of the quality of service which the identifier of a packet flow and this flow require are held at least.

[0033] It is desirable to use information required since the flow which requires a quality of service becomes identifiable as an identifier of a packet flow at a meaning. For example, in the cel to which an ATM communication network is transmitted, the identifier (a virtual path identifier (VPI), virtual channel identifiers (VCI), or these groups) which the virtual connection to whom this cel belongs holds is used as said identifier information. Moreover, in the IP packet flow according to IPv6 (Internet Protocol version 6), the group of the flow label currently described in the header of an IP packet and a source address is used as said identifier information.

[0034] As a demand quality-of-service item, the transfer rates (the maximum transfer rate, average transfer rate, etc.) of a packet, the transfer delay (the maximum transfer delay, average transfer delay, etc.) of a packet, a service level, etc. are raised. And a target quality of service is filled in as demand quality so that it may be made to correspond with the above-mentioned demand quality-of-service item.

[0035] Here, a service level is explained. A service level is the Predictive service (Predictive Quality of Service) or Controlled about which it argues in IETF (Internet Engineering Task Force). It is one of the quality-of-service items specified in Delay service (Controlled Delay Quality of Service), and extent of a service level is set up as a logical numeric value. In current and two

kinds of above-mentioned qualities of service, three kinds of service level values are specified (2 level 1, 3), and delay is lessened so relatively that a service level value is small.

[0036] The queue length notified of the offer quality-of-service Management Department 12 by the queue Monitoring Department 13 to the packet flow of arbitration, The quality of service which the packet belonging to this packet flow has covered based on the weight value notified from the weight value control section 14 is computed. It judges whether it is what satisfies the demand quality of service this quality of service is indicated to be by the demand quality-of-service attaching part 11, and when it is judged that it is necessary to improve a quality of service (it deteriorates depending on the case), that is notified to the weight value control section 14.

[0037] If the packet flow in which the current transfer rate value is less than the demand transfer rate value exists when the transfer rate is specified as a demand quality-of-service item, it will notify to the weight value control section 14 that the transfer rate of the packet belonging to this packet flow is made to increase. If the packet flow to which the current transfer delay value has exceeded the demand transfer delay value exists when transfer delay is specified as a demand quality-of-service item, it will notify to the weight value control section 14 that the transfer delay of the packet belonging to this flow is decreased. When the service level is specified as a demand quality-of-service item, it notifies to the weight value control section 14 so that it compares the transfer delay value covered between each packet flow, and the size relation of a transfer delay value may be made to follow the size relation of a service level, if the group of the packet flow the size relation of a transfer delay value and whose size relation of a service level do not correspond exists. In addition, the method which performs the notice of modification for every fixed period of a certain or the always operated method to the weight value control section 14 based on the comparison with the quality of service and demand quality of service which each packet flow has covered, or a comparison result can be considered.

[0038] The queue Monitoring Department 13 recognizes the number of the waiting packets for a transfer for every flow by supervising the buffer sections 151-15n, and notifies the offer quality-of-service Monitoring Department 12 of it. The weight value control section 14 holds the weight value currently assigned to each packet, notifies the offer quality-of-service Monitoring Department 12 in this weight value, and changes a weight value according to the instruction from this offer quality-of-service Monitoring Department 12.

[0039] Next, the modification approach of the weight value used with this operation gestalt and this weight value is explained. This "weight value" is dynamically changed based on the quality of service of each flow covered now, and enables it to offer more certainly the quality of service demanded from each flow by this.

[0040] Here, it shall write w_i [the weight value assigned to the packet flow i (the number of packet flows to which, as for $1 \leq i \leq n$, the transfer processing section 16 gives its service)], and the packet belonging to Flow i shall be accumulated in buffer 15i.

[0041] The weight value w_i shall express the rate of transmitting the packet to which the transfer processing section 16 belongs to Flow i to all the packets accumulated into 151-15n of buffer sections. That is, the weight value w_i shall fulfill the following conditions.

$0 \leq w_i \leq 1$ and $w_1 + w_2 + \dots + w_n \leq 1$ — the method which writes the weight value w_i in addition as a band which Flow i can obtain by the transfer processing section 16 is also considered.

[0042] Moreover, below, the quality of service Q_i covered to the demand quality of service R_i now writes $Q_i < R_i$ [that it is not satisfactory]. Moreover, when the quality of service Q_k covered now has received service still better than the demand quality of service R_k , it is written as $Q_k > R_k$.

[0043] An example of operation when the transfer processing section 16 performs a packet transfer to drawing 3 according to the above-mentioned "weight value" is shown. It is referred to as $n=4$, $w_1=0.5$, $w_2=0.25$, and $w_3=w_4=0.125$ in drawing 3. When the example of arrival of a packet shown in drawing 3 was followed and transfer spacing of the packet belonging to a flow 1 is a , transfer spacing of the packet to which transfer spacing of the packet belonging to a flow 2 belongs to $2a$, a flow 3, or a flow 4 is set to $4a$.

[0044] A flow chart shows actuation of the packet transfer section in case absolute quality, such as a transfer rate, and max / average transfer delay, is used for drawing 4 as a quality of service. In this packet transfer section, the quality of service Q_i first covered now to all the flows that are performing the call forwarding service in this packet transfer section is computed (step S51).

[0045] A quality of service Q_i is carried out based on the number L_i of are recording packets in each queue obtained at the weight value w_i and the queue Monitoring Department 13 which hold by the weight value control section 14, and is computed at the offer quality-of-service Monitoring Department 12. For example, a quality of service is computed according to the following count.

Transfer rate =(cel transfer rate in the transfer processing section) x of Flow i (weight value w_i)

Transfer delay =(number L_i of are recording packets) / of Flow i (transfer rate of Flow i)

Next, the quality of service Q_i computed as mentioned above is compared with the demand quality of service R_i currently recorded by the demand quality-of-service attaching part 11 at the offer quality-of-service Monitoring Department 12 (step S52). At this time, when the flow i the quality of service Q_i covered now is not satisfactory exists to the demand quality of service R_i , reallocation of (step S53) and the weight value w_i is performed (step S54).

[0046] Here, two kinds of flows of the flow k which becomes (1) $Q_j < R_{jj}$ [flow / becoming] (2) $Q_k > R_k$ can be considered as an example of the flow set as the object of the reallocation of a weight value.

[0047] The weight value w_j of the flow which belongs above (1) is made to increase as a plan of the reallocation of a weight value, and it is thought that the plan which decreases the weight value w_k of the flow which belongs above (2) is desirable. The network resource currently assigned to the flow (flow which belongs above (2)) which has received service superfluously by this will be assigned to the flow (flow which belongs above (1)) with which demand quality is not filled, and improvement in the quality of service which the flow which belongs above (1) wears can be expected.

[0048] The technique of setting the augend a of the weight value w_j of the flow which belongs above (1) as homogeneity to said all flows as the modification approach of the weight value in that case, and setting the decrement b of the weight value w_k of the flow which belongs above (2) as homogeneity to said all flows can be considered. In addition, when the number of flows which belongs the number of flows which belongs above (1) to n_1 and the above (2) is set to n_2 , it is desirable to set up a value and b value in which $axn_1 = bxn_2$ is materialized.

[0049] Or the technique of setting up the decrement of the weight value w_k of the flow k which belongs the augend of the weight value w_j of the flow j which belongs above (1) to $cx (R_j - Q_j)$ and the above (2) with $dx (Q_k - R_k)$ is also considered. In this case, $c \sum a_j = (R_j - Q_j)$ $d \sum a_k (Q_k - R_k)$

It is desirable to set up c value and d value which *****. Here, it is $\sum a_j$. About the total about the flow to which the amount of network resources is made to increase, it is $\sum a_k$. The total about a flow which decreases the amount of network resources is expressed.

[0050] If the reallocation of a weight value is completed according to the above-mentioned procedure, the comparison with the quality of service Q_i and the demand quality of service R_i which have actually covered the fixed passage of time at waiting (step S55) and its time, and the reallocation of a required weight value will be repeated, and will be performed.

[0051] In addition, as for allocation of the weight value w_i over each flow i at the time of packet transfer initiation, it is desirable to be carried out so that relation equivalent to the size relation of the quality of service which this flow requires may be materialized also to a weight value.

[0052] You may make it except from the object of the reallocation of a weight value in the above control about the flow whose difference of the demand quality of service R_i and the quality of service Q_i covered now is below the value specified beforehand.

[0053] Next, a flow chart shows actuation of the packet transfer section in case relative quality like a service level as a quality of service is used for drawing 5 . In this packet transfer section, the quality of service Q_i covered now to all the flows that are performing the call forwarding service in this packet transfer section is computed (step S61).

[0054] Next, the minimum value Q_{lmin} and Maximum Q_{lmax} of the quality of service Q_i computed as mentioned above to each service level l of every are calculated (step S62). At this time, relation called $Q_{lmin} < Q(l+1)_{max}$ is filled to a service level l and $l+1$ (the worst quality of service in a service level l). The service level l inferior to the best quality of service in a service level $l+1$. At the offer quality-of-service Monitoring Department 12, it supervises whether $l+1$ exists (step S63), and reallocation of the weight value w_i over the flow belonging to these service levels is performed to the service level l of which the above-mentioned relation consists, and $l+1$ (step S64).

[0055] To the flow which belongs, for example to a service level l as a plan of the reallocation of a weight value, the weight value w_s of the flow s which wears the quality of service which is less than $Q(l+1)_{max}$ is made to increase, and it is thought to the flow belonging to a service level $l+1$ that the plan which decreases the weight value w_t of the flow t which wears the quality of service exceeding Q_{lmin} is desirable. this modification plan — $Q_{lmin} > Q(l+1)_{max}$ — modification which forms the relation of the right quality of service which should be realized between relation l , i.e., a service level, and $l+1$ of dominance is aimed at.

[0056] The inside of the flow which belongs to a service level l as the modification approach of the weight value in that case, A quality of service sets the augend alpha of the weight value w_s of the flow (the number of flows is set to $n(l)$) which is less than $Q(l+1)_{max}$ as homogeneity to said all flows. The technique by which a quality of service sets the decrement beta of the weight value w_t of the flow (the number of flows is set to $n(l+1)$) exceeding Q_{lmin} as homogeneity to said all flows among the flows belonging to a service level $l+1$ can be considered. In addition, $\alpha_{n(l)} = \beta_{n(l+1)}$

It is desirable to set up alpha value and a beta value which *****.

[0057] Or when the quality of service which the flow s in which a quality of service is less than $Q(l+1)_{max}$ among the flows belonging to a service level l has covered is set to Q_s , The augend of the weight value w_s of Flow s gamma ($Q(l+1)_{max} - Q_s$), When the quality of service which the flow t to which a quality of service exceeds Q_{lmin} among the flows belonging to a service level $l+1$ has covered is set to Q_t , the technique of setting up the decrement of the weight value w_t of Flow t with delta ($Q_t - Q_{lmin}$) is also considered. In this case, $\gamma_{s} = (Q(l+1)_{max} - Q_s)$ $\delta_{t} = (Q_t - Q_{lmin})$

It is desirable to set up gamma which ***** and delta value. Here, it is γ_s . About the total about the flow to which the amount of network resources is made to increase, it is δ_t . The total about a flow which decreases the amount of network resources is expressed.

[0058] If the reallocation of a weight value is completed according to the above procedures, a comparison-related [between waiting (step S65) and the service level in the time / of dominance] and the reallocation of a required weight value will be repeated, and the fixed passage of time will be performed.

[0059] In addition, as for allocation of the weight value w_i over each flow i at the time of packet transfer initiation, between the flows which require a different service level, it is desirable to be carried out so that the direction of the weight value over the flow which requires a service level with more severe quality may become large.

[0060] In order to offer the packet transmittal mode in case the relative quality shown here is used, in each service level, a means to hold and update the size relation of the quality of service between the flows which require this level is needed. For example, a means to prepare the conversion table of a flow and a quality of service as shown in drawing 6 in the offer quality-of-service Monitoring Department 12 for every service level can be considered. Said conversion table sets in order and holds the flow identifier F_n belonging to a certain service level, and the quality of service Q_n which this flow has covered now according to the size relation of Q_n . In the example of drawing 6, a configuration as which the information about the flow to which the information about the flow which has covered the best quality has covered a rank 1 and quality good for the n -th is written down in Rank n is adopted.

[0061] this conversion table — every fixed period — or it is updated based on the quality of service which each flow in the time has covered whenever change arises in the number of packets of buffer circles which holds the packet belonging to the flow written down in this

conversion table. For example, as shown in drawing 6, the case where the quality of service of Flow Fn changes into Qn[from Qn]' is considered. At this time, only the size relation between Qn' after modification, and quality-of-service Qn-1 of the flow written down in the rank of those upper and lower sides and Qn+1 is compared. When the size relation of the quality of service which was contradictory in the rank on that occasion is realized, the location as which the information about the corresponding flow is filled in is replaced. The above-mentioned procedure is repeated, and if the size relation of the quality of service contradictory to a rank stops realizing, renewal of this table will be ended. In the example of drawing 6, since $Qn' > Qn-1$ is materialized, it is shown that compared size relation with Q' and the information about Flow Fn was finally written down in Rank m in the quality of service of the flow which changes to Rank n so that the information concerning Flow Fn in the information about flow Fn-1 may be written down in a rank n-1, and is continuously written down in the rank n-2.

[0062] In addition, it is desirable to prepare the function which samples the information about the flow specified from the conversion table belonging to the service level of arbitration, and the function which inserts the flow information which the above sampled into the conversion table belonging to the service level after modification at the offer quality-of-service Monitoring Department 12 supposing the case where the flow which requires modification to a different service level arises.

[0063] Since the extract of the maximum of the quality of service worn with this service level by using the conversion table shown in drawing 6 and the minimum value is obtained by extracting respectively the quality of service currently described at the rank 1 and Rank N (N expresses the number of flows belonging to this service level), it is easy. Moreover, since it can carry out by repeating the comparison with the quality of service about the flow of an adjoining rank, and performing it, more nearly high-speed updating also of size-related updating of the quality of service between flows is attained.

[0064] In addition, even if a configuration in case absolute quality, such as a transfer rate, and max / average transfer delay, is used as a quality of service explained while referring to drawing 4, and a configuration in case relative quality like a service level as a quality of service explained while referring to drawing 5 is used can be carried out independent and it uses together, it can carry out.

[0065] Although the operation gestalt of the packet transfer section which assigns a weight value for every flow was explained above, about some flows, these can be treated as "a group of a flow", and this invention can be applied also to the packet transfer section which assigns a weight value for every group of a flow or a flow, and controls a packet transfer.

[0066] Drawing 8 is an example of the packet transfer section which treats all the flows that require the same service level as one "a group of a flow", assigns a weight value to every group [of "flow"] about these flows, and was made to control the packet transfer.

[0067] Actuation of each part of the packet transfer section of drawing 8 is the same as that of the part to which drawing 1 corresponds fundamentally. In drawing 8, the buffer sections 151 and 15m of the point which holds one packet flow, respectively and assigns a network resource to each flow are the same as that of the operation gestalt of drawing 1. On the other hand, the buffer section 152 has held [not all] the flows as which the buffer section 154 requires a service level 3 for all the flows as which the buffer section 153 requires a service level 2 for all the flows that require a service level 1, respectively, and assigns a network resource for every flow to these flows, but assigns a network resource for every group of a flow. Therefore, the procedure of drawing 4 or drawing 5 is performed [flows / these] about the group of a flow.

[0068] And as mentioned above, the quality of service as which the group of each flow or a flow requires the "weight value" which controls the amount of network resources assigned to the group of each flow or a flow from the group of all flows or flows by making it change dynamically based on the difference of the quality of service which the group of each flow or a flow required, and the actually covered quality of service can be offered certainly. This invention is not limited to the gestalt of operation mentioned above, in the technical range, can deform variously and can be carried out.

[0069]

[Effect of the Invention] It becomes realizable [a packet transfer to which a demand quality of service is satisfied certainly] by according to this invention, supervising the quality of service which the group of each flow or a flow has covered, and changing dynamically the amount of network resources assigned to the group of each flow or a flow so that the difference of the quality of service and demand quality of service which have been covered now may be decreased, when the group of the flow with which are not satisfied of demand service, or a flow exists.

[0070] Moreover, it becomes realizable [a packet transfer to which the demanded service level is satisfied certainly] by assigning a network resource with which the relation of dominance between service levels and the relation of the actually covered quality of service of dominance become equivalent dynamically also to the group of the packet flow which requires a relative quality of service like a service level, or a flow according to this invention.

[Translation done.]

* NOTICES *

JPO and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] Drawing showing an example of the packet transfer section concerning the gestalt of operation of this invention

[Drawing 2] Drawing showing an example of the information held in a demand quality-of-service attaching part

[Drawing 3] Drawing showing the example of the packet transfer using a weight value of operation

[Drawing 4] The flow chart which shows the monitor of a quality of service, and the modification procedure of a weight value

[Drawing 5] The monitor of a quality of service, other flow charts which show the modification procedure of a weight value

[Drawing 6] Drawing showing an example of the conversion table of the flow and quality of service which are prepared for every service level

[Drawing 7] Drawing showing an example of operation of the conventional packet transmittal mode

[Drawing 8] Drawing showing other examples of the packet transfer section concerning the gestalt of operation of this invention

[Description of Notations]

11 -- Demand quality-of-service attaching part

12 -- Offer quality-of-service Monitoring Department

13 -- Queue Monitoring Department

14 -- Weight value control section

151, 152, 153, 154, 15n, 15m -- Buffer section

16 -- Transfer processing section

[Translation done.]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-238142

(43) 公開日 平成9年(1997)9月9日

(51) Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
H 0 4 L 12/28		9466-5K	H 0 4 L 11/20	G
12/46			H 0 4 Q 3/00	
12/56			H 0 4 L 11/00	3 1 0 C
H 0 4 Q 3/00		9466-5K	11/20	1 0 2 C

審査請求 未請求 請求項の数 9 O L (全 11 頁)

(21) 出願番号 特願平8-43542

(22) 出願日 平成8年(1996)2月29日

(71) 出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72) 発明者 大和 克己

神奈川県川崎市幸区小向東芝町1番地 株

式会社東芝研究開発センター内

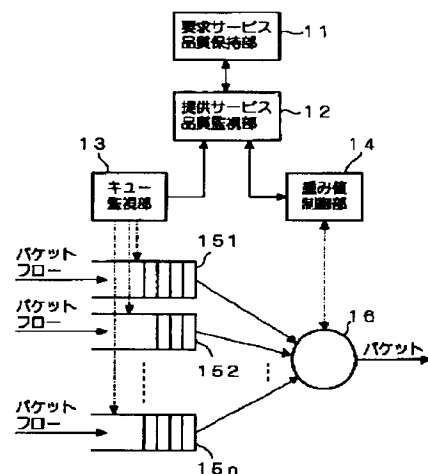
(74) 代理人 弁理士 鈴江 武彦

(54) 【発明の名称】 網リソース割り当て方法及び装置

(57) 【要約】

【課題】 全フロー又はフローの組に対し各フローまたはフローの組が要求するサービス品質を確実に提供できる網リソース割り当て方法を提供すること。

【解決手段】 パケット転送装置にて確保される網リソースを転送パケット内のパケット識別子を基に分類されるフロー又はフローの組毎に論理的に割り当て、パケットを転送する際には、該パケットの属するフロー又はフローの組に対し割り当てられた網リソース量のみを使用する方法であって、全フロー又はフローの組について現在割り当てられている網リソース量により該フロー又はフローの組が被るサービス品質と該フロー又はフローの組に予め要求されたサービス品質との間の差分を求め、この差分を全フロー又はフローの組において減少させるように各フロー又はフローの組に対する網リソース量の割り当てを変更させることを特徴とする。



【特許請求の範囲】

【請求項 1】通信情報が細分化された一包みの単位であるパケットを転送するパケット転送装置にて網リソースの割り当てを行う網リソース割り当て方法であって、パケット転送のために提供可能な網リソースを、各パケット内に記されているパケット識別子に基づき分類されるフローまたはフローの組毎に、該フローまたはフローの組に属するパケットを転送するために使用する所定量づつ論理的に割り当て、

各フローまたはフローの組について、現在割り当てられて 10 いる網リソース量により該フローまたはフローの組が被るサービス品質と、該フローまたはフローの組があらかじめ要求しているサービス品質との間の差分を求め、この差分を各フローまたはフローの組において減少させるように、各フローまたはフローの組に対する網リソース量の割り当てを変更することを特徴とする網リソース割り当て方法。

【請求項 2】前記各フローまたはフローの組に対する網リソース量の変更は、現在被っているサービス品質が 20 あらかじめ要求しているサービス品質を満足しないフローまたはフローの組に対しては網リソース量を増加し、現在被っているサービス品質があらかじめ要求しているサービス品質を満足するフローまたはフローの組に対しては網リソース量を減少するものであることを特徴とする請求項 1 に記載の網リソース割り当て方法。

【請求項 3】網リソース量を増加させるフローまたはフローの組に対する増加量を該フローまたはフローの組全てにおいて均一値 a とし、網リソース量を減少させるフローまたはフローの組に対する減少量を該フローまたは 30 フローの組全てにおいて均一値 b とし、網リソースを増加させるフローまたはフローの組の数を m_1 とし、網リソースを減少させるフローまたはフローの組の数を m_2 とした場合に、前記均一値 a および前記均一値 b を、 $a \times m_1 = b \times m_2$ が成立するように選択することを特徴とする請求項 2 に記載の網リソース割り当て方法。

【請求項 4】現在割り当てられている網リソース量により前記フローまたはフローの組 x が被るサービス品質を Q_x とし、前記フローまたはフローの組 x があらかじめ要求しているサービス品質を R_x とし、網リソース量を増加させるフローまたはフローの組 j に対する増加量を 40 $c \times (R_j - Q_j)$ とし (c は均一値)、網リソース量を減少させるフローまたはフローの組 k に対する減少量を $d \times (Q_k - R_k)$ とし (d は均一値)、前記網リソース量を増加させるフローまたはフローの組についての総和を Σ_j とし、前記網リソース量を減少させるフローまたはフローの組についての総和を Σ_k とした場合に、前記均一値 c および前記均一値 d を、 $c \times \Sigma_j (R_j - Q_j) = d \times \Sigma_k (Q_k - R_k)$ が成立するように選択することを特徴とする請求項 2 に記載の網リソース割り当て方法。

【請求項 5】通信情報が細分化された一包みの単位であるパケットを転送するパケット転送装置にて網リソースの割り当てを行う網リソース割り当て方法であって、パケット転送のために提供可能な網リソースを、各パケット内に記されているパケット識別子に基づき分類されるフローまたはフローの組毎に、該フローまたはフローの組に属するパケットを転送するために使用する所定量づつ論理的に割り当て、

サービス品質を相対的な優劣を示す情報にて要求している各フローまたはフローの組のうち、現在割り当てられて いる網リソース量により被るサービス品質の優劣関係と、あらかじめ要求しているサービス品質の優劣関係との間に相違が生じているものが存在する場合には、あらかじめ要求しているサービス品質の優劣関係が成り立つように、該当するフローまたはフローの組に対する網リソース量の割り当てを変更することを特徴とする網リソース割り当て方法。

【請求項 6】前記サービス品質としては、該サービス品質を要求する前記フローまたはフローの組が被るべきサービスのレベルを論理的に示す値が用いられるものであり、前記各フローまたはフローの組に対する網リソース量の変更は、各レベルに属する各フローまたはフローの組について現在割り当てられている網リソース量により被るサービス品質を求め、互いに異なるレベルに属する各フローまたはフローの組において、より優れたサービス品質を要求する第一のレベルに属するフローまたはフローの組が被るサービス品質のうち少なくとも一つの品質が、該第一のレベルよりも劣ったサービス品質を要求する第二のレベルに属するフローまたはフローの組のいずれかが被るサービス品質に比べ劣ったものである場合は、該第一のレベルに属するフローまたはフローの組に対し割り当てられている網リソース量を増加し、該第二のレベルに属するフローまたはフローの組に対し割り当てられている網リソース量を減少するものであることを特徴とする請求項 5 に記載の網リソース割り当て方法。

【請求項 7】網リソース量を増加させる前記第一のレベルに属するフローまたはフローの組に対する増加量を該フローまたはフローの組全てにおいて均一値 α とし、網リソース量を減少させる前記第二のレベルに属するフローまたはフローの組に対する減少量を該フローまたはフローの組全てにおいて均一値 β とし、網リソースを増加させるフローまたはフローの組の数を n_1 とし、網リソースを減少させるフローまたはフローの組の数を n_2 とした場合に、前記均一値 α および前記均一値 β を、 $\alpha \times n_1 = \beta \times n_2$ が成立するように選択することを特徴とする請求項 6 に記載の網リソース割り当て方法。

【請求項 8】現在割り当てられている網リソース量により前記フローまたはフローの組 x が被るサービス品質を Q_x とし、前記第一のレベルに属するフローまたはフロ

一の組において被るサービス品質のうち最悪のサービス品質を Q_{1min} とし、前記第二のレベルに属するフローまたはフローの組において被るサービス品質のうち最良のサービス品質を Q_{2max} とし、網リソース量を増加させる前記第一のレベルに属するフローまたはフローの組 s に対する増加量を $\gamma \times (Q_{2max} - Q_s)$ とし (γ は均一値)、網リソース量を減少させる前記第二のレベルに属するフローまたはフローの組 t に対する減少量を $\delta \times (Q_t - Q_{1min})$ とし (δ は均一値)、網リソース量を増加させるフローまたはフローの組についての総和を Σ とし、網リソース量を減少させるフローまたはフローの組についての総和を Σ_t とした場合、前記均一値 γ および前記均一値 δ を、
 $\gamma \times \Sigma, (Q_{2max} - Q_s) = \delta \times \Sigma_t, (Q_t - Q_{1min})$

が成立するように選択することを特徴とする請求項 6 に記載の網リソース割り当て方法。

【請求項 9】通信情報が細分化された一包みの単位であるパケットの転送のために提供可能な網リソースを、各パケット内に記されているパケット識別子に基づき分類されるフローまたはフローの組毎に、該フローまたはフローの組に属するパケットを転送するために使用する所定量づつ論理的に割り当てる手段と、

各フローまたはフローの組があらかじめ要求しているサービス品質もしくはその相対的な優劣関係を保持する手段と、

該フローまたはフローの組について、現在割り当てられている網リソース量により該フローまたはフローの組が被るサービス品質もしくはその相対的な優劣関係が前記保持する手段の保持内容を満足するものであるか否かを監視する手段とを備え、

前記割り当てる手段は、前記監視する手段の監視結果に基づいて網リソース量の割り当てを変更するものであることを特徴とする網リソース割り当て装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、パケット転送装置にて確保される網リソースをフロー単位またはフローの組単位に論理的に割り当てる網リソース割り当て方法及び装置に関する。

【0002】

【従来の技術】ATM (非同期転送モード; Asynchronous Transfer Mode) 通信網におけるセルと呼ばれる固定長のパケット、またはインターネットを用いた通信を行う際に用いられる IP (Internet Protocol) レベルにおいて定義されるパケット (可変長) 等の転送に際して、該パケットを送受信するユーザより、パケット転送に対するサービス品質の要求がなされた場合、通信網はこれらの要求品質を満足するようなパケットの転送を提供することが望まれる。ここで、サービス品質とは、例

えば任意のパケットが被る転送速度、最大/平均転送遅延といった絶対的な品質、あるいはサービスレベルと呼ばれる他のパケットが被る品質と比較した際の相対的な品質を表す。

【0003】これらの要求サービス品質を提供するために通信網において必須となるパケット転送のスケジューリングアルゴリズムがいくつか検討されている。例えば、到着時間の古いパケットから転送する FIFO (First In First Out) と呼ばれるポリシーに基づく待ち行列モデルに従ったパケットバッファを保持するパケット転送部において、該パケット転送部がサービスを行うパケットの属するフロー (場合によってはフローの組) 毎にあらかじめ優先権を割り当てておき、到着時間に関係なく優先権の高いフローに属するパケットから転送するというパケット転送方式がある (例えば、L. Kleinrock: "Queueing Systems, Volume II", Wiley-Interscience社)。また、各フローに対して重みづけを行い、パケット転送部に到着した各パケットは、該パケットが属するフローに対して割り当てられた重みに比例して被る転送サービス頻度が決定されるという WFQ (Weighted Fair Queueing) 方式も考えられている (例えば、D. D. Clark、他: "Supporting Real-Time Applications in an Integrated Services Packet Network: Architecture and Mechanism", ACM SIGCOMM '92)。WFQ 方式では、上記重みに相当する帯域が各フローに固定的に割り当てられるため、該帯域にて提供可能なサービスを各フローに属するパケットに対して行うことが可能である。

【0004】

【発明が解決しようとする課題】上述した優先権に基づくパケット転送方式では、優先権の低いフロー (あるいはフローの組) に対しては、パケット転送の機会は稀にしか与えられず、その結果、優先権の低いフローに対する品質保証を行うための制御を上記優先権制御に加えて提供する必要があると考えられる。また、上記の WFQ に基づくパケット転送方式では、各フローに割り当てる重みに従い被るサービス品質が当初予期していたサービス品質と異なる可能性もあり、その結果、このような事象に柔軟に対処することにより、要求サービス品質を常に満たすような動的な制御が必要であると考えられる。

【0005】本発明は、上記事情を考慮してなされたものであり、全てのフローまたはフローの組に対して、各フローまたはフローの組が要求するサービス品質を確実に提供できるようにするための網リソース割り当て方法及び装置を提供することを目的とする。

【0006】

【課題を解決するための手段】本発明 (請求項 1) は、通信情報が細分化された一包みの単位であるパケット (例えば、ATMセル、IPパケット) を転送するパケット転送装置にて網リソース (例えば、パケット転送速

度)の割り当てを行う網リソース割り当て方法であつて、パケット転送のために提供可能な網リソースを、各パケット内に記されているパケット識別子に基づき分類されるフローまたはフローの組毎に、該フローまたはフローの組に属するパケットを転送するために使用する所定量づつ論理的に割り当て、各フローまたはフローの組について、現在割り当てられている網リソース量により該フローまたはフローの組(x)が被るサービス品質(Qx)と、該フローまたはフローの組(x)があらかじめ要求しているサービス品質(Rx)との間の差分を求め、この差分(|Rx-Qx|)を各フローまたはフローの組において減少させるように、各フローまたはフローの組に対する網リソース量の割り当てを変更することを特徴とする。

【0007】上記の各フローまたはフローの組に対する網リソース量の割り当ての変更は、動的に、例えば所定のタイミングで繰返し行うことが望ましい。パケット識別子は、例えば、ATM通信方式におけるVPI/VC I(仮想パス/チャネル識別子)、あるいはIPプロトコル(バージョンIPv6)におけるフローラベルとソースアドレスの組等を用いる。

【0008】フローの組とは、例えば、同一のサービス品質を要求する複数のフローを、1つの組として取り纏めたものであり、この場合、フローの組に対して網リソースが割り当てられる。

【0009】本発明によれば、サービス品質として転送速度あるいは最大/平均転送遅延等の絶対的な品質を用いる各フローまたはフローの組に対し、夫々が要求するサービス品質を確実に満足させるようなパケット転送を提供することが可能となる。

【0010】本発明(請求項2)は、請求項1において、前記各フローまたはフローの組に対する網リソース量の変更は、現在被っているサービス品質があらかじめ要求しているサービス品質を満足しないフローまたはフローの組に対しては網リソース量を増加し、現在被っているサービス品質があらかじめ要求しているサービス品質を満足するフローまたはフローの組に対しては網リソース量を減少するものであることを特徴とする。

【0011】本発明(請求項3)は、請求項2において、網リソース量を増加させるフローまたはフローの組に対する増加量を該フローまたはフローの組全てにおいて均一値aとし、網リソース量を減少させるフローまたはフローの組に対する減少量を該フローまたはフローの組全てにおいて均一値bとし、網リソースを増加させるフローまたはフローの組の数をm1とし、網リソースを減少させるフローまたはフローの組の数をm2とした場合に、前記均一値aおよび前記均一値bを、 $a \times m1 = b \times m2$ が成立するように選択することを特徴とする。

【0012】本発明(請求項4)は、請求項2において、現在割り当てられている網リソース量により前記フ

ローまたはフローの組xが被るサービス品質をQxとし、前記フローまたはフローの組xがあらかじめ要求しているサービス品質をRxとし、網リソース量を増加させるフローまたはフローの組jに対する増加量を $c \times (Rj - Qj)$ とし(cは均一値)、網リソース量を減少させるフローまたはフローの組kに対する減少量を $d \times (Qk - Rk)$ とし(dは均一値)、前記網リソース量を増加させるフローまたはフローの組についての総和を Σ_j とし、前記網リソース量を減少させるフローまたはフローの組についての総和を Σ_k とした場合に、前記均一値cおよび前記均一値dを、 $c \times \Sigma_j = d \times \Sigma_k$ が成立するように選択することを特徴とする。

【0013】本発明(請求項5)は、通信情報が細分化された一包みの単位であるパケットを転送するパケット転送装置にて網リソースの割り当てを行う網リソース割り当て方法であつて、パケット転送のために提供可能な網リソースを、各パケット内に記されているパケット識別子に基づき分類されるフローまたはフローの組毎に、該フローまたはフローの組に属するパケットを転送するために使用する所定量づつ論理的に割り当て、サービス品質を相対的な優劣を示す情報にて要求している各フローまたはフローの組のうち、現在割り当てられている網リソース量により被るサービス品質の優劣関係と、あらかじめ要求しているサービス品質の優劣関係との間に相違が生じているものが存在する場合には、あらかじめ要求しているサービス品質の優劣関係が成り立つように、該当するフローまたはフローの組に対する網リソース量の割り当てを変更することを特徴とする。

【0014】上記の各フローまたはフローの組に対する網リソース量の割り当ての変更は、動的に、例えば所定のタイミングで繰返し行うことが望ましい。本発明によれば、サービスレベルのように相対的なサービス品質を要求する各フローまたはフローの組に対し、夫々が要求するサービス品質を確実に満足させるようなパケット転送を提供することが可能となる。

【0015】本発明(請求項6)は、請求項5において、前記サービス品質としては、該サービス品質を要求する前記フローまたはフローの組が被るべきサービスのレベルを論理的に示す値が用いられるものであり、前記各フローまたはフローの組に対する網リソース量の変更は、各レベルに属する各フローまたはフローの組について現在割り当てられている網リソース量により被るサービス品質を求め、互いに異なるレベルに属する各フローまたはフローの組において、より優れたサービス品質を要求する第一のレベルに属するフローまたはフローの組が被るサービス品質のうち少なくとも一つの品質が、該第一のレベルよりも劣ったサービス品質を要求する第二のレベルに属するフローまたはフローの組のいずれかが被るサービス品質に比べ劣ったものである場合は、該第

一のレベルに属するフローまたはフローの組に対し割り当てられている網リソース量を増加し、該第二のレベルに属するフローまたはフローの組に対し割り当てられている網リソース量を減少するものであることを特徴とする。

【0016】本発明（請求項7）は、請求項6において、網リソース量を増加させる前記第一のレベルに属するフローまたはフローの組に対する増加量を該フローまたはフローの組全てにおいて均一値 α とし、網リソース量を減少させる前記第二のレベルに属するフローまたはフローの組に対する減少量を該フローまたはフローの組全てにおいて均一値 β とし、網リソースを増加させるフローまたはフローの組の数を n_1 とし、網リソースを減少させるフローまたはフローの組の数を n_2 とした場合に、前記均一値 α および前記均一値 β を、 $\alpha \times n_1 = \beta \times n_2$ が成立するように選択することを特徴とする。

【0017】本発明（請求項8）は、請求項6において、現在割り当てられている網リソース量により前記フローまたはフローの組 x が被るサービス品質を Q_x とし、前記第一のレベルに属するフローまたはフローの組において被るサービス品質のうち最悪のサービス品質を Q_{1min} とし、前記第二のレベルに属するフローまたはフローの組において被るサービス品質のうち最良のサービス品質を Q_{2max} とし、網リソース量を増加させる前記第一のレベルに属するフローまたはフローの組 s に対する増加量を $\gamma \times (Q_{2max} - Q_s)$ とし（ γ は均一値）、網リソース量を減少させる前記第二のレベルに属するフローまたはフローの組 t に対する減少量を $\delta \times (Q_t - Q_{1min})$ とし（ δ は均一値）、網リソース量を増加させるフローまたはフローの組についての総和を Σ とし、網リソース量を減少させるフローまたはフローの組についての総和を Σ とした場合、前記均一値 γ および前記均一値 δ を、 $\gamma \times \Sigma$ 、 $(Q_{2max} - Q_s) = \delta \times \Sigma$ 、 $(Q_t - Q_{1min})$ が成立するように選択することを特徴とする。

【0018】本発明（請求項9）は、通信情報が細分化された一包みの単位であるパケットの転送のために提供可能な網リソースを、各パケット内に記されているパケット識別子に基づき分類されるフローまたはフローの組毎に、該フローまたはフローの組に属するパケットを転送するために使用する所定量づつ論理的に割り当てる手段と、各フローまたはフローの組があらかじめ要求しているサービス品質もしくはその相対的な優劣関係を持する手段と、該フローまたはフローの組について、現在割り当てられている網リソース量により該フローまたはフローの組が被るサービス品質もしくはその相対的な優劣関係が前記保持する手段の保持内容を満足するものであるか否かを監視する手段とを備え、前記割り当てる手段は、前記監視する手段の監視結果に基づいて網リソース量の割り当てを変更するものであることを特徴とする

網リソース割り当て装置。

【0019】上記の各フローまたはフローの組に対する網リソース量の割り当ての変更は、動的に、例えば所定のタイミングで繰返し行うことが望ましい。本発明によれば、サービス品質として転送速度あるいは最大／平均転送遅延等の絶対的な品質を用いる各フローまたはフローの組およびサービスレベルのように相対的なサービス品質を要求する各フローまたはフローの組に対し、夫々が要求するサービス品質を確実に満足させるようなパケット転送を提供することが可能となる。

【0020】

【発明の実施の形態】以下、図面を参照しながら発明の実施の形態を説明する。本実施形態に係るパケット転送装置内に設けられるパケット転送部は、例えば図1に示すような構成を有する。複数のパケットフロー送出のスケジューリングを行うパケット転送部では、到着したパケットを該パケットの属するフロー毎に、論理的あるいは物理的に異なる待ち行列（キュー）151～15n内にて一時保持し、転送処理部16では、各フローに対して割り当てられている「重み」をもとに、転送すべきパケットの属するフロー（待ち行列）を選択し、後に送出を行なう。

【0021】ここで、パケットとは、通信情報を転送する際に該情報を細分化した一包みの単位であり、例えばATM通信方式におけるセルと呼ばれる固定長パケットや、TCP、IP等のプロトコルにおいてフォーマット等があらかじめ定められている可変長パケットなどを示す。また、フローとは、前記パケット内に記されているパケット識別子が同一であるパケットの集合、または前もって規定した複数のパケット識別子のうちいずれかの識別子を持つパケットの集合を表し、ATM通信方式ではVPI/VCI（仮想パス／チャネル識別子）、またIPパケットの場合、バージョンがIPv6であれば、フローラベルとソースアドレスの両値により、パケット識別子を一意に規定する。

【0022】転送処理部16では、各フローに与えられる網リソース量（パケット転送速度）が、あらかじめフローに対して割り当てられた重みの値に比例する値をとるよう、複数のフロー間でのパケット転送に関するスケジューリングを行う。なお、このスケジューリング機能は、例えばWFQ方式を採用することにより実現できる。

【0023】本実施形態にて用いる重み値としては、相当するフローに割り当てられる網リソース量を採用する方式、あるいは該フローが出力回線を使用できる割合を採用する方式などが考えられる。前者の場合、重み値の総和は出力回線が提供できる網リソース量となり、後者の場合、重み値の総和は1となる。

【0024】本実施形態のパケット転送部は、各フローに割り当てられる網リソース量を制御する重み値を動的

に変更する重み値制御部 14 を保持する。これは、各フロー毎の蓄積パケット数を監視するキュー監視部 13 内の情報と現在の重み値を元に、提供サービス品質監視部 12 において全てのフローに対して満足のいくサービスがなされているかを監視し、必要に応じて、任意のフローの重み値を動的に変更する。例えば、各フローが要求したパケット転送遅延、パケット転送速度等の要求サービス品質が満足に提供されていない場合、あるいはサービスレベルの優劣関係と、提供されているサービス品質の優劣関係が異なるフローの組が存在する場合に、重み値を変更する。

【0025】なお、詳しくは後述する様に、幾つかのフロー、例えば同一サービスレベルを要求する全てのフローについてはこれらを「フローの組」として扱い、フローの組毎に重み値を割り当ててパケット転送の制御を行うこともできる。

【0026】以下、本発明の実施形態をさらに詳細に説明する。なお、以下では、各フロー毎に重み値を割り当てるものとして説明する。図 1 に、本実施形態に係るパケット転送装置内に設けられるパケット転送部を示す。

【0027】本パケット転送部は、要求サービス品質保持部 11、提供サービス品質監視部 12、キュー監視部 13、重み値制御部 14、バッファ部 151~15n、転送処理部 16 を備えている。

【0028】要求サービス品質保持部 11 は、各パケットフローに対する要求サービス品質を保持しておく。提供サービス品質監視部 12 は、本パケット転送部において各パケットフローに対して提供しているサービス品質を算出し、該サービス品質を要求サービス品質保持部 11 内にて保持されている要求サービス品質と比較する。

【0029】キュー監視部 13 は、各パケットフロー毎に蓄積パケット数を監視する。重み値制御部 14 は、各パケットフローに対して割り当てる転送速度（帯域）を決定する「重み値」を、提供サービス品質監視部 12 の指示に従い変更する。

【0030】バッファ部 151~15n は、各々のパケットフローを収容する。転送処理部 16 は、バッファ部 151~15n 内の任意のパケットを抽出して転送する。

【0031】ここで従来においては図 7 に示すように、各フローからのサービス品質の要求を受けた要求サービス品質保持部 11 において、各フローに対して割り当てる重み値を固定的に決定するか、または要求サービス品質にもとづいた優先制御を行っている。これに対して、本実施形態では、この重み値を、各フローが要求したサービス品質と実際に被っているサービス品質との差分に基づき動的に変化させ、全てのフローに対し各フローが要求するサービス品質を確実に提供できるようにしている。

【0032】以下、本パケット転送部の構成および動作

についてさらに詳しく説明する。要求サービス品質保持部 11 は、例えば図 2 に示すような情報を保持している。要求サービス品質保持部 11 における保持情報として、パケットフローの識別子、該フローが要求するサービス品質の項目、および要求品質が少なくとも保持される。

【0033】パケットフローの識別子としては、サービス品質を要求するフローが一意に識別可能となるために必要な情報が用いられることが望ましい。例えば、ATM 通信網を転送されるセルにおいては、該セルが属する仮想コネクションが保持する識別子（仮想パス識別子（VPI）、仮想チャネル識別子（VCI）、あるいはこれらの組）が前記識別子情報として用いられる。また、IPv6（インターネットプロトコルバージョン 6）に従った IP パケットフローにおいては、IP パケットのヘッダ内に記されているフローラベルとソースアドレスの組が前記識別子情報として用いられる。

【0034】要求サービス品質項目としては、パケットの転送速度（最大転送速度、平均転送速度、等）、パケットの転送遅延（最大転送遅延、平均転送遅延、等）、サービスレベル等があげられる。そして、上記の要求サービス品質項目と対応させるように、目標とすべきサービス品質が要求品質として記入される。

【0035】ここで、サービスレベルに関して説明する。サービスレベルは、IETF（Internet Engineering Task Force）にて議論されている Predictive サービス（Predictive Quality of Service）または Controlled Delay サービス（Controlled Delay Quality of Service）において規定されているサービス品質項目の一つであり、論理的数値としてサービスレベルの程度が設定される。現在、上記の 2 種類のサービス品質においては、3 種類のサービスレベル値が規定されており（レベル 1、2、3）、サービスレベル値が小さいほど相対的に遅延が少ないとされている。

【0036】提供サービス品質管理部 12 は、任意のパケットフローに対して、キュー監視部 13 より通知されるキュー長、重み値制御部 14 より通知される重み値を元に、該パケットフローに属するパケットが被っているサービス品質を算出し、該サービス品質が要求サービス品質保持部 11 に記載されている要求サービス品質を満足するものであるか否かを判断し、サービス品質を改善（場合によっては改悪）する必要があると判断した場合は、重み値制御部 14 へその旨を通知する。

【0037】要求サービス品質項目として転送速度が指定されている場合、現在の転送速度値が要求転送速度値を下回っているパケットフローが存在すれば、該パケットフローに属するパケットの転送速度を増加させるよう重み値制御部 14 へ通知する。要求サービス品質項目として転送遅延が指定されている場合、現在の転送遅延値

10

20

30

40

50

が要求転送遅延値を上回っているパケットフローが存在すれば、該フローに属するパケットの転送遅延を減少させるよう重み値制御部14へ通知する。要求サービス品質項目としてサービスレベルが指定されている場合、各パケットフロー間で被っている転送遅延値を比較し、転送遅延値の大小関係とサービスレベルの大小関係とが一致しないパケットフローの組が存在すれば、転送遅延値の大小関係をサービスレベルの大小関係に従わせるよう、重み値制御部14へ通知する。なお、各パケットフローが被っているサービス品質と要求サービス品質との比較、または比較結果に基づいた重み値制御部14への変更の通知は、ある一定周期毎に行う方式、あるいは常に動作させる方式が考えられる。

【0038】キュー監視部13は、フロー毎の転送待ちパケット数を、バッファ部151～15nを監視することにより認識し、提供サービス品質監視部12へ通知する。重み値制御部14は、各々のパケットに対して割り当てられている重み値を保持しており、該重み値を提供サービス品質監視部12へ通知を行い、該提供サービス品質監視部12からの命令に従い、重み値を変更する。

【0039】次に、本実施形態にて用いる重み値と、該重み値の変更方法に関して説明する。この「重み値」は、現在被っている各フローのサービス品質に基づいて動的に変化させるものであり、これによって各フローから要求されたサービス品質をより確実に提供できるようにしている。

【0040】ここで、パケットフロー i ($1 \leq i \leq n$; n は転送処理部16がサービスを行うパケットフロー数)に割り当てられた重み値を w_i と表記し、フロー i に属するパケットはバッファ15iに蓄積されるものとする。

【0041】重み値 w_i は、バッファ部151～15n内に蓄積されている全パケットに対して転送処理部16がフロー i に属するパケットの転送を行う割合を表すものとする。つまり、重み値 w_i は、以下の条件を満たすものとする。

$$0 \leq w_i \leq 1 \quad \text{かつ} \quad w_1 + w_2 + \dots + w_n \leq 1$$

なお、重み値 w_i を、転送処理部16によりフロー i が得ることのできる帯域として表記する方式も考えられる。

【0042】また、以下では、要求サービス品質 R_i に対して現在被っているサービス品質 Q_i が満足のいかないことを、 $Q_i < R_i$ と表記する。また、現在被っているサービス品質 Q_k が要求サービス品質 R_k よりもさらに良質のサービスを受けている場合、 $Q_k > R_k$ と表記する。

【0043】図3に、上記の「重み値」に従って転送処理部16がパケット転送を行った場合の動作例を示す。

図3では、 $n=4$ 、 $w_1=0.5$ 、 $w_2=0.25$ 、 $w_3=w_4=0.125$ としている。図3に示したパケッ

トの到着例に従うと、フロー1に属するパケットの転送間隔が a であった場合、フロー2に属するパケットの転送間隔は $2a$ 、フロー3やフロー4に属するパケットの転送間隔は $4a$ となる。

【0044】図4に、サービス品質として転送速度、最大／平均転送遅延等の絶対的な品質が用いられる場合のパケット転送部の動作をフローチャートにて示す。本パケット転送部では、まず、該パケット転送部にて転送サービスを行っている全てのフローに対して、現在被っているサービス品質 Q_i を算出する(ステップS51)。

【0045】サービス品質 Q_i は、重み値制御部14にて保持している重み値 w_i およびキュー監視部13にて得られている各キュー内の蓄積パケット数 L_i を元にして、提供サービス品質監視部12において算出する。例えば、以下のような計算に従い、サービス品質を算出する。

$$\text{フロー } i \text{ の転送速度} = (\text{転送処理部でのセル転送速度}) \times (\text{重み値 } w_i)$$

$$\text{フロー } i \text{ の転送遅延} = (\text{蓄積パケット数 } L_i) \div (\text{フロー } i \text{ の転送速度})$$

次に、上記のようにして算出したサービス品質 Q_i と、要求サービス品質保持部11にて記録されている要求サービス品質 R_i とを、提供サービス品質監視部12にて比較する(ステップS52)。このとき、要求サービス品質 R_i に対して、現在被っているサービス品質 Q_i が満足のいかないフロー i が存在する場合は(ステップS53)、重み値 w_i の再配分を行う(ステップS54)。

【0046】ここで、重み値の再配分の対象となるフローの一例として、

(1) $Q_j < R_j$ なるフロー j

(2) $Q_k > R_k$ なるフロー k

の2種類のフローが考えられる。

【0047】重み値の再配分の方針としては、上記

(1)に属するフローの重み値 w_j を増加させ、上記

(2)に属するフローの重み値 w_k を減少させる方針が望ましいと考えられる。これにより、過剰にサービスを受けているフロー(上記(2)に属するフロー)へ割り当てられている網リソースを、要求品質を満たしていないフロー(上記(1)に属するフロー)へ割り当てることとなり、上記(1)に属するフローが被るサービス品質の向上が期待できる。

【0048】その際の重み値の変更方法としては、上記(1)に属するフローの重み値 w_j の増加量 a を前記フロー全てに対して均一に設定し、上記(2)に属するフローの重み値 w_k の減少量 b を前記フロー全てに対して均一に設定する手法が考えられる。なお、上記(1)に属するフロー数を n_1 、上記(2)に属するフロー数を

$$a \times n_1 = b \times n_2$$

が成立するような a 値および b 値を設定することが望ましい。

【0049】あるいは、上記(1)に属するフロー j の重み値 w_j の増加量を $c \times (R_j - Q_j)$ 、上記(2)に属するフロー k の重み値 w_k の減少量を $d \times (Q_k - R_k)$ と設定する手法も考えられる。この場合、 $c \sum_j (R_j - Q_j) = d \sum_k (Q_k - R_k)$ が成立するような c 値および d 値を設定することが望ましい。ここで、 \sum_j は、網リソース量を増加させるフローについての総和を、 \sum_k は、網リソース量を減少させるフローについての総和を表している。

【0050】上記の手順に従って重み値の再配分が終了すれば、一定時間の経過を待ち(ステップS55)、その時点で実際に被っているサービス品質 Q_i と要求サービス品質 R_i との比較および必要な重み値の再配分を繰り返す。

【0051】なお、パケット転送開始時における、各フロー i に対する重み値 w_i の配分は、該フローが要求するサービス品質の大小関係と等価な関係が重み値に対しても成立するように行なわれるのが望ましい。

【0052】以上の制御において、要求サービス品質 R_i と現在被っているサービス品質 Q_i との差分が予め規定した値以下であるフローについては、重み値の再配分の対象から除外するようにしても良い。

【0053】次に、図5に、サービス品質としてサービスレベルのような相対的な品質が用いられる場合のパケット転送部の動作をフローチャートにて示す。本パケット転送部では、該パケット転送部にて転送サービスを行っている全てのフローに対して、現在被っているサービス品質 Q_i を算出する(ステップS61)。

【0054】次に、各サービスレベル l 毎に、上記のようにして算出したサービス品質 Q_i の最小値 Q_{lmin} および最大値 Q_{lmax} を求める(ステップS62)。このとき、サービスレベル l 、 $l+1$ に対して、 $Q_{lmin} < Q_{(l+1)max}$

という関係を満たす(サービスレベル l における最悪のサービス品質が、サービスレベル $l+1$ における最良のサービス品質より劣っている)サービスレベル l 、 $l+1$ が存在するか否かを、提供サービス品質監視部12において監視し(ステップS63)、上記関係が成り立つサービスレベル l 、 $l+1$ に対しては、これらのサービスレベルに属するフローに対する重み値 w_i の再配分を行う(ステップS64)。

【0055】重み値の再配分の方針としては、例えばサービスレベル l に属するフローに対しては、 $Q_{(l+1)max}$ を下回るサービス品質を被るフロー s の重み値 w_s を増加させ、サービスレベル $l+1$ に属するフローに対しては、 Q_{lmin} を上回るサービス品質を被るフロー t の重み値 w_t を減少させる方針が望ましいと考えられる。この変更方針は、

$$Q_{lmin} > Q_{(l+1)max}$$

なる関係、すなわちサービスレベル l 、 $l+1$ の間にて成り立つべき正しいサービス品質の優劣関係を成立させるような変更を目指したものである。

【0056】その際の重み値の変更方法としては、サービスレベル l に属するフローのうち、サービス品質が $Q_{(l+1)max}$ を下回るフロー(フロー数を $n(l)$ とする)の重み値 w_s の増加量 α を前記フロー全てに対して均一に設定し、サービスレベル $l+1$ に属するフローのうち、サービス品質が Q_{lmin} を上回るフロー(フロー数を $n(l+1)$ とする)の重み値 w_t の減少量 β を前記フロー全てに対して均一に設定する手法が考えられる。なお、

$$\alpha \times n(l) = \beta \times n(l+1)$$

が成立するような α 値および β 値を設定することが望ましい。

【0057】あるいは、サービスレベル l に属するフローのうち、サービス品質が $Q_{(l+1)max}$ を下回るフロー s が被っているサービス品質を Q_s とした場合、フロー s の重み値 w_s の増加量を $\gamma(Q_{(l+1)max} - Q_s)$ 、サービスレベル $l+1$ に属するフローのうち、サービス品質が Q_{lmin} を上回るフロー t が被っているサービス品質を Q_t とした場合、フロー t の重み値 w_t の減少量を $\delta(Q_t - Q_{lmin})$ と設定する手法も考えられる。この場合、 $\gamma \sum_s (Q_{(l+1)max} - Q_s) = \delta \sum_t (Q_t - Q_{lmin})$

が成立するような γ 、 δ 値を設定することが望ましい。ここで、 \sum_s は、網リソース量を増加させるフローについての総和を、 \sum_t は、網リソース量を減少させるフローについての総和を表している。

【0058】上記のような手順に従って重み値の再配分が終了すれば、一定時間の経過を待ち(ステップS65)、その時点でのサービスレベル間での優劣関係の比較および必要な重み値の再配分を繰り返す。

【0059】なお、パケット転送開始時における、各フロー i に対する重み値 w_i の配分は、異なるサービスレベルを要求するフロー間では、より品質の厳しいサービスレベルを要求するフローに対する重み値の方が大きくなるように行なわれるのが望ましい。

【0060】ここで示した相対的な品質が用いられる場合のパケット転送方式を提供するためには、各々のサービスレベルにおいて、該レベルを要求するフロー間でのサービス品質の大小関係を保持および更新する手段が必要となる。例えば、提供サービス品質監視部12にて、図6に示すようなフローとサービス品質との対応表を各サービスレベル毎に用意する手段が考えられる。前記対応表は、あるサービスレベルに属するフロー識別子 F_n と、該フローが現在被っているサービス品質 Q_n を、 Q_n の大小関係に従い順序づけを行い、保持してい

る。図6の例では、最も良い品質を被っているフローに関する情報がランク1に、そしてn番目に良い品質を被っているフローに関する情報がランクnに記入されるような構成を採用している。

【0061】本対応表は、一定周期毎にあるいは本対応表に記入されているフローに属するパケットを収容するバッファ部内のパケット数に変化が生じる度に、その時点での各フローが被っているサービス品質をもとに更新される。例えば、図6に示すように、フローF_nのサービス品質がQ_nからQ_n'に変更した場合を考える。このとき、変更後のQ_n'と、その上下のランクに記入されているフローのサービス品質Q_{n-1}、Q_{n+1}との大小関係のみを比較する。その際に、ランクに矛盾したサービス品質の大小関係が成り立っている場合には、該当するフローに関する情報を記入する位置を入れ替える。上記手順を繰り返し、ランクに矛盾したサービス品質の大小関係が成り立たなくなれば、本表の更新を終了する。図6の例では、Q_n' > Q_{n-1}が成立するので、ランクnにはフローF_{n-1}に関する情報を、ランクn-1にはフローF_nに関する情報を記入するように入れ替え、続けてランクn-2に記入されているフローのサービス品質をQ'との大小関係を比較し、最終的にはフローF_nに関する情報はランクmに記入されたことを示している。

【0062】なお、異なるサービスレベルへの変更を要求するフローが生じた場合を想定して、任意のサービスレベルに属する対応表より指定されたフローに関する情報を抜き取る機能、そして変更後のサービスレベルに属する対応表内へ前記の抜き取ったフロー情報を挿入する機能が提供サービス品質監視部12にて用意されることが望ましい。

【0063】図6に示した対応表を用いることにより、該サービスレベルにて被るサービス品質の最大値、最小値の抽出が、ランク1、ランクN(Nは該サービスレベルに属するフロー数を表す)に記されているサービス品質を各々抽出することにより得られるため、容易である。また、フロー間のサービス品質の大小関係の更新も、隣接するランクのフローに関するサービス品質との比較を繰り返し行うことにより行えるので、より高速な更新が可能となる。

【0064】なお、図4を参照しながら説明したサービス品質として転送速度、最大/平均転送遅延等の絶対的な品質が用いられる場合の構成と、図5に参照しながら説明したサービス品質としてサービスレベルのような相対的な品質が用いられる場合の構成とは、独立実施可能であり、また併用しても実施可能である。

【0065】以上では、各フロー毎に重み値を割り当てるパケット転送部の実施形態について説明したが、幾つかのフローについてはこれらを「フローの組」として扱い、フローまたはフローの組毎に重み値を割り当ててパ

ケット転送の制御を行うパケット転送部にも、本発明は適用可能である。

【0066】図8は、同一サービスレベルを要求する全てのフローを一つの「フローの組」として扱い、これらのフローについては「フローの組」毎に重み値を割り当ててパケット転送の制御を行うようにしたパケット転送部の一例である。

【0067】図8のパケット転送部の各部分の動作は、基本的には図1の対応する部分と同様である。図8において、バッファ部151、15mは、それぞれ1つのパケットフローを収容し、各フローに網リソースを割り当てる点は、図1の実施形態と同様である。これに対して、バッファ部152はサービスレベル1を要求する全てのフローを、バッファ部153はサービスレベル2を要求する全てのフローを、バッファ部154はサービスレベル3を要求する全てのフローをそれぞれ収容しており、これらのフローに対しては、フロー毎に網リソースを割り当てるのではなく、フローの組毎に網リソースを割り当てる。従って、これらのフローについては、図4や図5の手順は、フローの組について行われる。

【0068】そして前述したように、各フローまたはフローの組に割り当てられる網リソース量を制御する「重み値」を、各フローまたはフローの組が要求したサービス品質と実際に被っているサービス品質との差分に基づき動的に変化させることにより、全てのフローまたはフローの組に対し各フローまたはフローの組が要求するサービス品質を確実に提供することができる。本発明は、上述した実施の形態に限定されるものではなく、その技術的範囲において種々変形して実施することができる。

【0069】

【発明の効果】本発明によれば、各フローまたはフローの組が被っているサービス品質を監視し、要求サービスを満足しないフローまたはフローの組が存在する場合は、現在被っているサービス品質と要求サービス品質との差分を減少させるように各フローまたはフローの組に割り当てる網リソース量を動的に変化させることにより、要求サービス品質を確実に満足させるようなパケット転送の実現が可能となる。

【0070】また、本発明によれば、サービスレベルのように相対的なサービス品質を要求するパケットフローまたはフローの組に対しても、サービスレベル間の優劣関係と実際に被っているサービス品質の優劣関係が等価となるような網リソースの割り当てを動的に行うことにより、要求されたサービスレベルを確実に満足させるようなパケット転送の実現が可能となる。

【図面の簡単な説明】

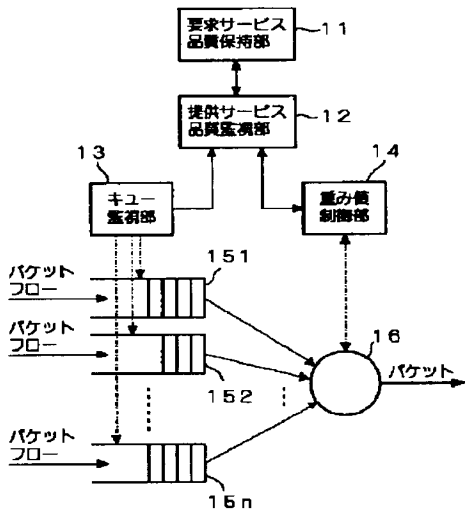
【図1】本発明の実施の形態に係るパケット転送部の一例を示す図

【図2】要求サービス品質保持部にて保持する情報の一例を示す図

【図 3】 重み値を用いたパケット転送の動作例を示す図
【図 4】 サービス品質の監視、重み値の変更手順を示すフローチャート
【図 5】 サービス品質の監視、重み値の変更手順を示す他のフローチャート
【図 6】 各サービスレベル毎に用意されるフローとサービス品質との対応表の一例を示す図
【図 7】 従来のパケット転送方式の実施の一例を示す図
【図 8】 本発明の実施の形態に係るパケット転送部の他*

* の例を示す図
【符号の説明】
1 1…要求サービス品質保持部
1 2…提供サービス品質監視部
1 3…キュー監視部
1 4…重み値制御部
1 5 1, 1 5 2, 1 5 3, 1 5 4, 1 5 n, 1 5 m…バッファ部
1 6…転送処理部

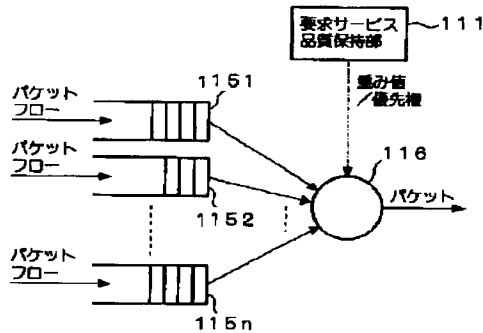
【図 1】



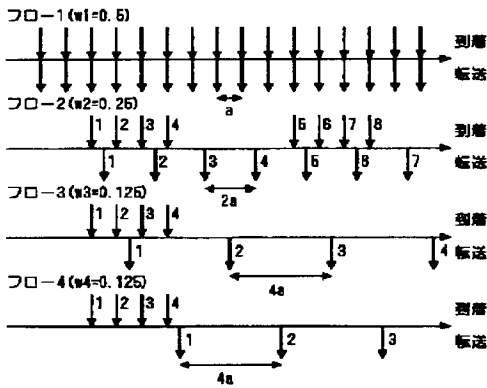
【図 2】

識別子	要求項目	要求品質
# 1	転送速度	10Mb/s
# 2	転送遅延	0.1ms
# 3	サービスレベル	レベル 1
# 4	転送速度	5Mb/s
# 5	サービスレベル	レベル 2

【図 7】



【図 3】

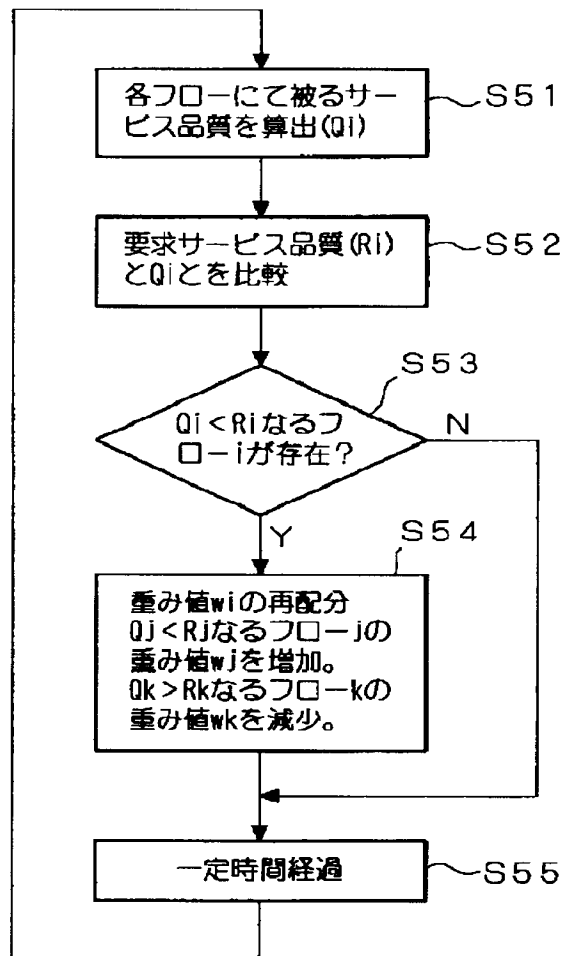


【図 6】

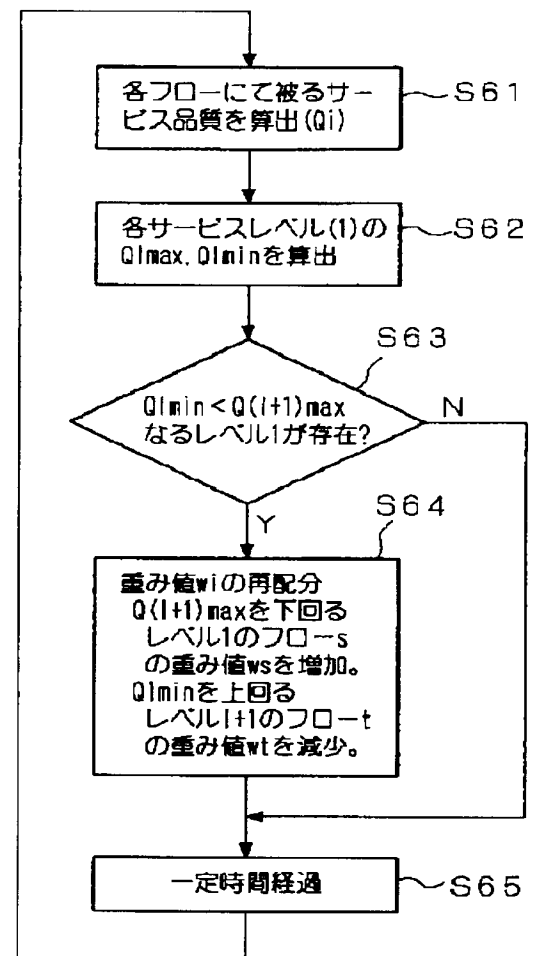
ランク	フロー	サービス品質
1	F1	Q1
2	F2	Q2
...
n-1	Fn-1	Qn-1
n	Fn	Qn-Qn'
n+1	Fn+1	Qn+1
...
N-1	FN-1	QN-1
N	FN	QN

ランク	フロー	サービス品質
1	F1	Q1
2	F2	Q2
...
n	Fn	Qn
n+1	Fn+1	Qn+1
...
N-1	FN-1	QN-1
N	FN	QN

【図4】



【図5】



【図8】

